



Análise do padrão de metilação do genoma na epilepsia mioclônica juvenil

B. S. Lopes¹, D.C.F. Bruno¹, W. Souza¹, J. C. Geraldis¹, B. S. Carvalho³, M.K.M. Alvim²,
C.L. Yasuda², F. Cendes², I. Lopes-Cendes¹

1 Department of Medical Genetics and Genomic Medicine, 2 Department of Neurology; School of Medical Sciences, 3 Department of Statistics, Institute of Mathematics, Statistics and Science Computing. University of Campinas (UNICAMP) and the Brazilian Institute of Neuroscience and Neurotechnology (BRAINN), Campinas, SP, Brazil.

Introdução

Estima-se que existam aproximadamente 50 milhões de pessoas com epilepsia no mundo [1]. A Epilepsia Mioclônica Juvenil (EMJ) é uma epilepsia genética generalizada, onde 50% dos casos tem histórico familiar positivo [15]. EMJ representa certa 10% de todos os casos de epilepsias. EMJ se manifesta inicialmente por volta da puberdade, entre 14-16 anos, no entanto são encontrados casos com uma variabilidade muito maior, entre os 8 a 26 anos [2].

Alguns estudos apontam alguns genes candidatos a suscetibilidade genética de EMJ, com alterações identificadas principalmente no gene *BRD2* e no receptor *GABRA1* [3,4]. Contudo, a maior parte dos pacientes com EMJ não apresentam variantes nos genes candidatos identificados, não existindo um consenso sobre possíveis biomarcadores para EMJ, sugerindo um desenvolvimento multifatorial [5].

Mecanismos epigenéticos são influenciados por fatores ambientais e participam da regulação gênica. Estes podem modificar a expressão gênica, alterar o fenótipo sem modificar a sequência de DNA. Existem 3 principais mecanismos epigenéticos: (1) metilação de DNA, (2) modificação de histonas e (3) ação de RNAs não codificantes. O mecanismo epigenético mais estudado é a metilação do DNA, e um de seus mecanismos ocorre através da adição de um grupamento metil no carbono 5 da molécula de citosina seguida de guanina, o dinucleotídeo CpG, estes em geral são clusterizados em ilhas CpG nas regiões promotoras de genes, o que comumente leva à inibição da transcrição [6]. Ainda, a regulação epigenética possui um importante papel em todos os aspectos da função neuronal, desde a embriogênese e desenvolvimento inicial do cérebro, até a expressão gênica específica de determinado tecido, ao silenciamento global de genes [6].

Fatores ambientais e epigenéticos podem contribuir na amplificação dos efeitos de variantes genéticas, podendo ter papel crucial no desenvolvimento de epilepsias [7]. Diversos estudos já propuseram o papel da epigenética na regulação, desenvolvimento e predisposição à diferentes tipos de epilepsia, incluindo a EMJ [8-10]. Portanto, é razoável acreditar que a desregulação epigenética possa ter um papel na predisposição e/ou participação no desenvolvimento de EMJ.

Em EMJ, dois trabalhos já sugeriram padrões de metilação do DNA para os genes *BRD2*, *NKCC1* e *KCC2* [8,9]. Tais achados, nos levaram a hipótese de que mecanismos epigenéticos, mais especificamente a metilação do DNA, possam estar envolvidos na etiologia e/ou predisposição da EMJ. A maior parte dos estudos que avaliaram a metilação em EMJ olharam apenas para metilações em alguns genes sabidamente relacionados [8,9]. Até o momento, para EMJ, não conhecemos nenhum estudo que investigou o padrão de metilação do DNA em todo o

genoma (metiloma) em pacientes. Por tanto, propomos este estudo como o primeiro que sequenciou e caracterizou o metiloma de pacientes com EMJ através da técnica de sequenciamento do genoma completo utilizando a conversão de bissulfito de sódio (WGBS).

Material e métodos

Nesse estudo foram avaliados 3 pacientes e 3 indivíduos controle saudáveis. Os dois grupos são similares em sexo e idade, média de 30 ± 8 anos. Os pacientes com EMJ foram selecionados levando em conta apenas o uso uma droga antiepiléptica - ácido valpróico. As amostras de DNA foram extraídas previamente através da coleta de sangue periférico em tubos com ácido etilenodiaminotetracético (EDTA). O DNA de cada amostra foi convertido com bissulfito de sódio (EZ DNA Methylation-Gold Kit by Zymo Research) e em seguida foi construída a biblioteca (Accel-NGS Methyl-Seq DNA Library Kit by Swift Biosciences). O sequenciamento do genoma completo tratado com bissulfito de sódio (WGBS) foi realizado em plataforma HiSeq-X. Os dados brutos provenientes do sequenciamento foram analisados, processados, trimados, deduplicados e alinhados com o genoma de referência (GRCh38) com o auxílio de ferramentas e softwares de bioinformática. Os níveis da metilação do DNA foram calculados com o auxílio do software Bismark. A identificação das DMRs foi feita com base no pipeline BSmooth no pacote bsseq [16]. O mapeamento e anotação das DMRs como sua localização gênica, condição de metilação, sobreposição com ilhas CpG, dentre outras informações foram obtidas utilizando os pacotes annotatr, biomaRt e genomation com auxílio do R. Ainda, foram utilizadas as ferramentas do UCSC Genome Browser e ENSEMBL. Algumas informações provenientes destas ferramentas foram chamadas por requisição API em Python.

Resultados e Discussão

Sequenciamento do DNA convertido com bissulfito de sódio mostra alta cobertura

O WGBS teve uma alta eficiência de conversão por bissulfito de sódio, em média 98.5% para todas as 6 amostras. Os dados brutos tiveram uma média de 952 ± 10.38 milhões de reads, com uma cobertura média de $45x \pm 0.49$. Após trimagem e remoção das sequências deduplicadas foi obtido cerca de 77% de alinhamentos de reads únicos, equivalente à uma média de 109.6 ± 14.9 milhões de reads, com média de cobertura citosinas de $31x \pm 1.35$ e cobertura média do genoma completo de $21x \pm 1.67$.

Diferentes regiões funcionais ao longo do genoma se sobrepõem às DMRs mapeadas

5.337 DMRs, com $p < 0,01$ foram identificadas sendo 2.967 hipermetiladas e 2.370 hipometiladas. Estas estão distribuídas em diferentes regiões no genoma, como pode ser observado no gráfico 1.

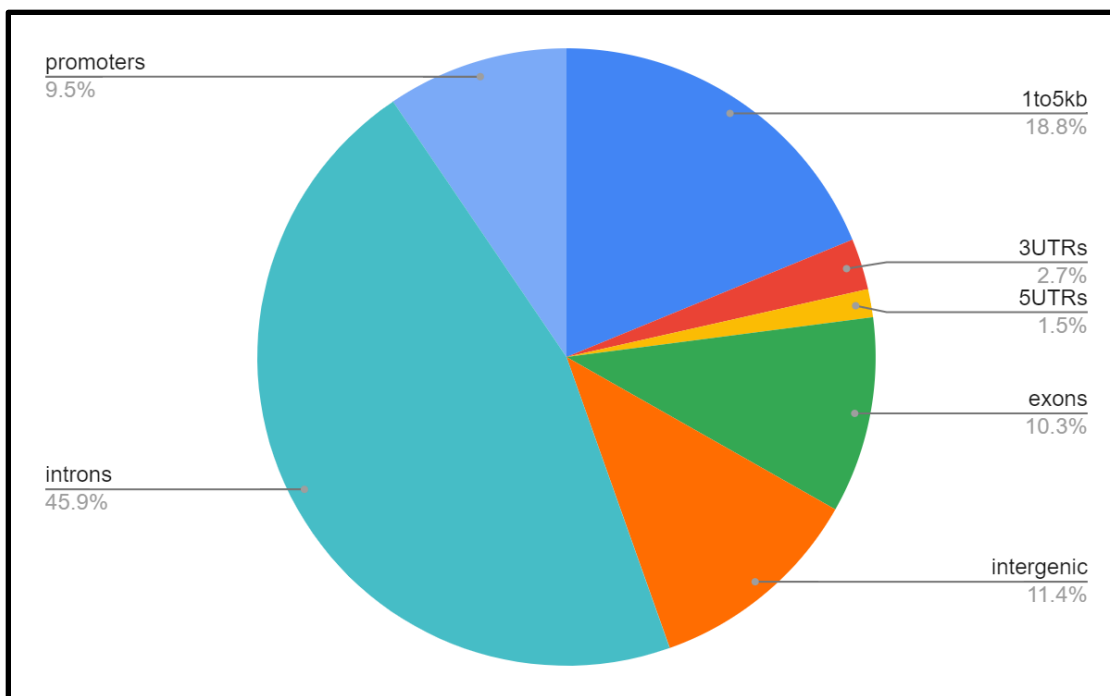


Gráfico 1. Distribuição das DMRs em relação à sua co-localização com regiões genômicas. Da esquerda para a direita- promoters: sequência de DNA onde se inicia a transcrição do gene pela ligação da proteína RNA polimerase; 1to5kb: regiões que se estendem de 1 até 5 kb a partir do sítio de início da transcrição; 3UTR: vai do códon de terminação até a região da cauda poli-A; 5UTRs: se inicia na extremidade ponta 5' e acaba no primeiro códon de início da tradução; exons: região do gene codante para amino ácidos; intergenic: regiões intergênicas, ou seja, regiões de DNA não-codante que está entre genes; Introns: região não-codante para amino ácidos.

Mapeando as DMRs, utilizando as coordenadas genômicas, vimos que estas se sobrepõem à 4.216 genes. É válido citar que cada uma das DMRs pode se sobrepôr à mais de um gene e mais de uma região genômica, além disso, os genes podem estar sobrepostos com mais de uma DMR. A anotação destes genes nos mostraram que 2.864 genes são codificadores de proteínas, 1801 são ncRNAs e 240 são pseudogenes. Dentre os ncRNAs temos outras categorias de menor número (miscRNAs, valrRNAs, TRCgene, entre outras).

Interessantemente, embora a metilação do DNA seja reconhecida por ocorrer predominantemente em ilhas CpG, nós identificamos que apenas 2,5% das DMRs se sobrepõem a estas regiões. Com isso, possivelmente o motivo pelo qual os estudos com metilação em EMJ existentes até o momento não tenham encontrado mais genes candidatos, é que estes focaram estudar regiões ricas em ilhas CpG. A exemplo trabalhos anteriores, como de Pathak S. et al. (2018), estudaram regiões enriquecidas em ilhas CpG de genes já relacionados com a EMJ, com tudo, nossos resultados nos levam a notar que doenças multifatoriais, como a EMJ, em geral são complexas e multigênicas [5], e que o presente estudo possui grande importância para a futura seleção de novos genes candidatos e o entendimento quanto ao papel destes no desenvolvimento da EMJ.

Tabela 1. Top 8 DMRs com metilação acima de 20% e sua sobreposição a genes

Localização da DMR	Metilação	ID Gene	ID Ensembl	Função molecular UniProt
2:231280843-231280998	hyper	<i>ARMC9</i>	ENSG00000135931	ND
2:239192933-239193100	hyper	<i>HDAC4</i>	ENSG00000068024	Ativação de fator de transcrição; histona deacetilase; ligação de proteínas; proteína deacetilase; regulação RNA polimerase II; co repressor da transcrição
16:12767863-12768596	hyper	<i>CPPED1</i>	ENSG00000103381	Atividade fosfatase fosfoproteína
20:63612882-63613023	hyper	<i>GMEB2</i>	ENSG00000101216	Atividade de ligação RNA polimerase II; atividade co-ativador da transcrição
3:71755844-71756021	hyper	<i>FOXP1</i>	ENSG00000114861	Sítio de ligação sequência-específica para promotor; Ligação de fator de transcrição; associação de proteínas com DNA
2:3316590-3316826	hyper	<i>EIPR1</i>	ENSG00000032389	ND
6:87407246-87407467	hyper	<i>CFAP206</i>	ENSG00000272514	ND
7:148398074-148398838	hyper	<i>CNTNAP2</i>	ENSG00000174469	Ligação de enzima

Observando a tabela 1, percebemos que os resultados são variados, quanto à sua função molecular. Estes genes também variam quanto à função biológica, enquanto alguns têm funções muito ligadas à regulação da expressão gênica, como o *GMEB2* e o *HDAC4*, outros regulam a resposta inflamatória, como o *CPPED1* e o *FOXP1*, existem também aqueles que codificam proteínas constituintes celulares, como o *ARMC9* e o *CFAP206*, alguns ainda participam da neurogênese, como o *CNTNAP2*, e até de vias de secreção hormonal, como o *EIPR1*.

É interessante observar que dos 8 genes aqui citados, sendo todos codificadores de proteínas, 5 já foram relacionados com epilepsias por outros estudos [10-14], reiterando que o estudo do padrão de metilação do genoma é importante para entendermos melhor a EMJ, uma vez que com isto reforçamos e trazemos maior significância para a relação destes genes com esta epilepsia, visto que estes foram relacionados em outros contextos, que não pelo estudo da diferença de metilação. Além disso, trouxemos aqui apenas 8 genes, porém considerando os mais de 4.200 genes encontrados e pelo menos 297 destes já foram relacionados com epilepsias, observamos

Por fim, concluímos que o metiloma de pacientes com EMJ apresenta padrões de regiões diferencialmente metiladas quando comparado à indivíduos saudáveis; e que o presente estudo mostra a importância da caracterização do metiloma dos pacientes com EMJ, uma vez que não olhamos somente para regiões enriquecidas em ilhas CpG, e sim observamos que 98% das DMRs estão localizados fora destas. EMJ é um distúrbio complexo e multigênico, portanto, este trabalho abre novos caminhos para estudos futuros que poderão melhor elucidar como os resultados aqui citados apresentam relação com o fenótipo da EMJ nos levando à sua melhor compreensão.

Bibliografia

1. Fisher, R. S., Acevedo. et al. (2014), ILAE Official Report: A practical clinical definition of epilepsy. *Epilepsia*, 55(4):475–482, 2014. doi: 10.1111/epi.12550

2. Kasteleijn- Nolst Trenité, Dorothée G. A., Bettina Schmitz, Dieter Janz, Antonio V. Delgado-Escueta, Pierre Thomas, Edouard Hirsch, Holger Lerche, et al. (2013). "Consensus on Diagnosis and Management of JME: From Founder's Observations to Current Trends." *Epilepsy & Behavior: E&B* 28 (July): S87–90.
3. Velíšek, Libor, Enyuan Shang, Jana Velíšková, Tamar Chachua, Stephania Macchiarulo, Giorgi Maglakelidze, Debra J. Wolgemuth, and David A. Greenberg. 2011. "GABAergic Neuron Deficit as an Idiopathic Generalized Epilepsy Mechanism: The Role of BRD2 Haploinsufficiency in Juvenile Myoclonic Epilepsy." *PloS One* 6 (8): e23656.
4. dos Santos, B. P. et al.(2017) Genetic susceptibility in Juvenile Myoclonic Epilepsy: Systematic review of genetic association studies. *PLoS One* 12.
5. A. Bird. (2007), Perceptions of epigenetics. *Nature*. 447. doi: 10.1038/nature05913
6. Millan M. J.,(2013), An epigenetic framework for neurodevelopmental disorders: From pathogenesis to potential therapy. *Neuropharmacology*, Volume 68, Pages 2-82. doi: 10.1016/j.neuropharm.2012.11.015.
7. A. Bird. (2007), Perceptions of epigenetics. *Nature*. 447. doi: 10.1038/nature05913
8. Pathak S., Miller J., et al. (2018), DNA methylation of the BRD2 promoter is associated with juvenile myoclonic epilepsy in Caucasians. *Epilepsia*. 59: 1011– 1019. doi: 10.1111/epi.14058 11.Genç, F., Kara, M., et al. (2019), Methylation of cation-chloride cotransporters NKCC1 and KCC2 in patients with juvenile myoclonic epilepsy. *Neurol Sci* . doi: 10.1007/s10072-019-03743-4.
9. Hong-Yu Long., Li Feng., et al. (2017) Blood DNA methylation pattern is altered in mesial temporal lobe epilepsy. *Scientific Reports* , volume 7, Article number: 43810. doi: 10.1038/srep43810
10. Shandra, Oleksii, Solomon L. Moshé, and Aristeia S. Galanopoulou. 2017. "Chapter Three - Inflammation in Epileptic Encephalopathies." In *Advances in Protein Chemistry and Structural Biology*, edited by Rossen Donev, 108:59–84. Academic Press.
11. Epi4K Consortium, Epilepsy Phenome/Genome Project, Andrew S. Allen, Samuel F. Berkovic, Patrick Cossette, Norman Delanty, Dennis Dlugos, et al. 2013. "De Novo Mutations in Epileptic Encephalopathies." *Nature* 501 (7466): 217–21.
12. Mullen, Saul A., Gemma L. Carvill, Susannah Bellows, Marta A. Bayly, Holger Trucks, Dennis Lal, Thoman Sander, et al. 2013. "Copy Number Variants Are Frequent in Genetic Generalized Epilepsy with Intellectual Disability." *Neurology* 81 (17): 1507–14.
13. Jay, Kristy, Amit Mitra, Taylor Harding, David Matthes, and Brian Van Ness. 2019. "Identification of a de Novo FOXP1 Mutation and Incidental Discovery of Inherited Genetic Variants Contributing to a Case of Autism Spectrum Disorder and Epilepsy." *Molecular Genetics & Genomic Medicine* 7 (7): e00751.
14. Peñagarikano, Olga, Brett S. Abrahams, Edward I. Herman, Kellen D. Winden, Amos Gdalyahu, Hongmei Dong, Lisa I. Sonnenblick, et al. 2011. "Absence of CNTNAP2 Leads to Epilepsy, Neuronal Migration Abnormalities, and Core Autism-Related Deficits." *Cell* 147 (1): 235–46.
15. Martínez-Juárez I. E., Alonso M. E., et al. (2006) Juvenile myoclonic epilepsy subsyndromes: family studies and long-term follow-up, *Brain*, Volume 129, Issue 5, Pages 1269–1280, <https://doi.org/10.1093/brain/awl048>
16. Hansen KD, Langmead B, Irizarry RA. (2012), BSmooth: from whole genome bisulfite sequencing reads to differentially methylated regions. *Genome Biol.* 2012;13(10):R83. doi:10.1186/gb-2012-13-10-r83